**Magazine**
R412

**Primer**

# Plant epigenetics

**Judith Bender**

## Introduction

Epigenetic changes in gene expression have fascinated scientists over several decades. These processes have received particular attention in plants, where they can result in beautiful variations in conspicuous phenotypes such as pigmentation. Epigenetic control is also a key issue in the development of transgenic plants with appropriate expression from newly introduced transgene segments.

The term ‘epigenetic’ refers to heritable gene expression patterns determined by how the DNA of a gene is packaged rather than its primary DNA sequence. Within tightly packed DNA, genes are not readily available to the transcription machinery and are poorly expressed. Normally the patterns of DNA packaging are carefully controlled to give predictable patterns of gene expression. However, the process can occasionally go awry to cause altered gene expression. This primer will focus on well characterized examples of epigenetic changes in plants that shed light on the mechanisms underlying this fundamental gene control process.

## Determinants of DNA packaging

In higher organisms, DNA is packaged into the nucleus of the cell by association with histone proteins; this DNA–protein complex is chromatin. Some regions of the genome are loosely packaged into euchromatin, whereas other regions are tightly packaged into heterochromatin. One factor that determines chromatin patterning is modification of histone proteins by attachment of small chemical groups to particular amino acid side chains. Specific patterns of histone modification are thought to recruit specific chromatin remodeling proteins that direct either heterochromatin or euchromatin formation.

In mammalian and plant genomes, chromatin patterning is also determined by the attachment of methyl groups to cytosine residues in the DNA by cytosine methyltransferases. When a region of genomic DNA has cytosine methylation it is typically assembled into heterochromatin. Methylated DNA appears to recruit methyl-DNA binding proteins, which in turn recruit histone-modifying enzymes and chromatin-remodeling factors necessary for heterochromatin formation. Cytosine methylation is a fundamental epigenetic mark that can be maintained after each round of DNA replication because the template strand of DNA will retain the modification. Although changes in the cytosine methylation mark often correlate with epigenetic variation, there are also likely to be cases where chromatin changes occur independently of methylation.

In mammals, DNA methylation marks are reprogrammed during early embryogenesis and altered methylation patterns are not usually transmitted to progeny. In plants, however, it seems that DNA methylation changes can persist throughout development and can be inherited between generations.

## Epigenetics in maize

Over the course of man’s domestication of maize, many strains with striking patterns of kernel or plant pigmentation have been selected for cultivation. These strains have provided a rich source of epigenetic variation in pigment gene expression. One such case that has been examined at the molecular level is the expression of a transcription factor gene that controls pigment synthesis, the *B* gene. The *B* gene is necessary for purple pigmentation of plant tissues. Several decades ago, an unusual behavior of a particular darkly pigmented *B* variant was observed: this purple strain would occasionally yield progeny that were green because they no longer fully expressed the *B* gene. More curiously, when the purple strain was crossed with a green variant, the hybrid plant and all its resulting progeny were green. Thus, the high expression state of *B* inherited from the purple parent was efficiently and permanently converted into the low expression state of the green parent by putting the two chromosomes together in the same nucleus. This phenomenon, called ‘paramutation’, seems to involve different epigenetic states of *B* gene regulatory sequences that result in different levels of transcription initiation.

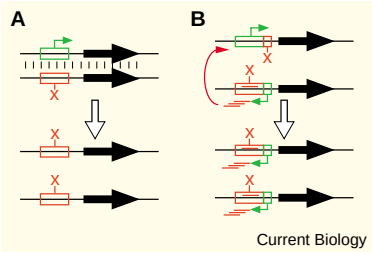


Figure 1. Mechanisms for paramutation. The *B* coding region is shown by a black arrow. Upstream regulatory sequences show euchromatin (green box) and heterochromatin (red box); expression (green arrow) and silencing (red X). (A) Pairing and direct transfer of chromatin components leads to silencing. (B) An aberrant RNA species (red lines) with identity to the *B* regulatory sequences triggers chromatin changes and silencing. Note that this model involves a competition between two opposing promoters in the upstream regulatory sequences.

*B* gene paramutation raises a number of challenges. Where is the exact control region that governs *B* transcription? What changes in chromatin structure or cytosine methylation might occur on the key sequences to convert the high expression state to a low expression state? And how is the epigenetic information from the low expression state chromosome efficiently communicated to the high expression state chromosome? One possibility is that the two chromosomes directly interact, and that chromatin components from the low expression chromosome are transferred to the high expression chromosome

(Figure 1). Alternatively, the low expression chromosome may produce a diffusible signal that is targeted to the high expression chromosome. An attractive candidate for this signal would be an RNA molecule with sequence identity to the DNA of the *B* expression control region. This is not an outlandish idea, as there is mounting evidence that, in plants, unusual RNA species can trigger cytosine methylation of related genomic DNA sequences. For example, infection of plants with RNA viruses can sometimes induce cytosine methylation of genomic DNA that has sequence identity to the viral genome.

Another approach to dissecting the mechanism of *B* paramutation is to characterize mutant maize strains defective for the process. One such mutant, called *mediator of paramutation1* (*mop1*), not only blocks *B* paramutation but also blocks epigenetic silencing of other pigment control genes. Clearly, epigenetic changes at the *B* locus are just one symptom of a key system for silencing gene expression. Cloning of the *MOP1* gene will thus identify a pivotal regulator of epigenetic states in maize.

**Epigenetics in Arabidopsis**

The materials for studying epigenetic variation in *Arabidopsis* come from two general sources: rearrangements in genome structure that trigger chromatin structure changes, and mutations in the cellular factors that regulate epigenetic patterning. A well-characterized system involving a naturally occurring gene rearrangement that leads to dramatic epigenetic alterations comes from studies of the *Arabidopsis PAI* genes. The *PAI* genes encode an enzyme necessary for synthesis of the amino acid tryptophan. In the majority of *Arabidopsis* isolates, PAI enzyme is encoded by two nearly identical genes located on two different chromosomes. In such strains, the *PAI* genes are stably expressed and lack cytosine methylation. However, in a minority of wild *Arabidopsis* isolates, one of the *PAI* loci is rearranged to carry an inverted repeat arrangement of two mirror-image *PAI* genes running into each other. Strikingly, in these unusual isolates, both the rearranged inverted repeat *PAI* genes and the outlying singlet *PAI* gene are densely covered with cytosine methylation over their regions of sequence identity. The *PAI* inverted repeat locus provides the signal for this methylation, because when it is combined with unmethylated *PAI* genes via genetic crosses with a 'normal' *Arabidopsis* strain, unmethylated *PAI* genes become densely methylated in the hybrid plants.

This ability of one locus to alter the epigenetic patterning of a related locus elsewhere in the genome is a variation on the theme of paramutation. However, a key difference between the *Arabidopsis PAI* case and the maize *B* case is that *PAI* involves an obvious change in DNA sequence as the trigger of the epigenetic change. Another difference is that the *PAI* inverted repeat locus can cause an epigenetic change at a *PAI* gene on a different chromosome. Whether the low expression *B* locus could signal to a high expression target *B* gene at a new chromosomal location is not known. Yet the general mechanisms proposed for *B* paramutation, direct interactions between the chromosomal loci, or a diffusible signal that moves from the trigger locus to the target locus, are also possible mechanisms for the *PAI* system.

As mentioned above, RNA with sequence identity to the affected genes is an attractive candidate for a diffusible signal that promotes epigenetic changes. Could the *PAI* inverted repeat be producing an unusual RNA product? One of the *PAI* genes in the inverted repeat is transcribed from a novel upstream regulatory sequence that lies beyond the methylated region. Such *PAI* methylated strains may be lucky to have this upstream sequence, as their remaining *PAI* genes are silenced by methylation of their normal proximal sequences. However, if transcription through the inverted repeat provides an RNA trigger for *PAI* methylation and silencing, then maybe they aren't so lucky after all!

In both *Arabidopsis* and tobacco, the regulatory sequences of a transgene can be efficiently methylated and silenced by a second trigger transgene carrying a transcribed inverted repeat of the regulatory sequences. This finding strengthens the idea that unusual RNAs expressed from an inverted repeat might promote epigenetic changes at homologous regions of the genome. This finding also suggests that any plant gene could be targeted for silencing with an appropriate transcribed regulatory sequence inverted repeat transgene.

To facilitate genetic screens for mutations that disrupt *PAI* epigenetic changes, a reporter strain has been created by mutation of the sole expressed *PAI* gene in the inverted repeat. This strain accumulates a blue fluorescent intermediate in the tryptophan pathway due to the epigenetic block in PAI enzyme levels (Figure 2). Reductions in the density of *PAI* gene methylation in this strain lead to proportional reductions in blue fluorescence. For example, mutations in either of two *Arabidopsis* cytosine methyltransferase genes, *MET1* or *CMT3*, lead to partial loss of *PAI* methylation and partial suppression of blue fluorescence. These experiments suggest that methylation density can act as a rheostat to adjust intermediate expression levels.

Another approach to understanding epigenetic control in *Arabidopsis* is to find loci whose expression changes when a component of the control machinery is mutated. A well-characterized mutant background for this type of analysis is deficient in cytosine methylation due to mutation of the *Decrease in DNA Methylation 1* (*DDM1*) gene. *DDM1* encodes a protein related to a yeast chromatin remodeling factor,

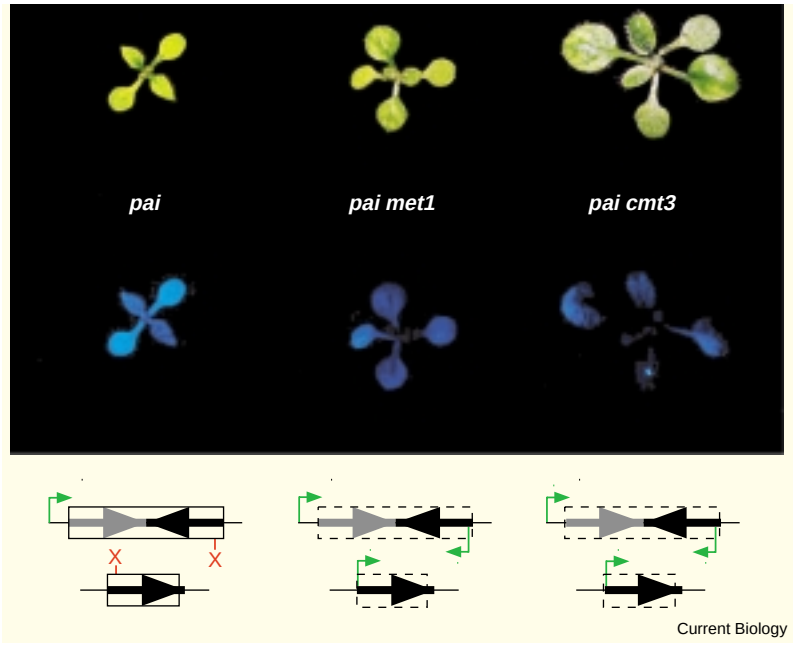


Figure 2. The *Arabidopsis pai* mutant strain provides a blue fluorescent reporter for *PAI* gene silencing.
The upper panel shows plants under visible light, the middle panel shows plants under ultra-violet light, and the lower panel shows diagrams of *PAI* gene methylation states in parental, *met1* mutant, or *cmt3* mutant backgrounds. *PAI* genes (thick black arrows); mutated genes (thick grey arrows); full and partial cytosine methylation (solid and dashed boxes); silencing (red X); and expression (green arrow) are shown.

SWI2/SNF2. Thus DDM1, like its mammalian homologue LSH1, might promote chromatin changes required for the normal action of cytosine methyltransferase enzymes at methylation target loci around the genome.

The original *ddm1* mutant isolates were morphologically normal. However, upon inbreeding they gave occasional abnormal progeny including dwarf plants, plants with a delay in flowering, and plants with aberrant flowers and fertility defects. Some of these variants correspond, either directly or indirectly, to changes in epigenetic control at sensitive sites in the genome. The *ddm1* mutant can thus be thought of as an 'epimutator' background that destabilizes epigenetic programming.

A somewhat paradoxical type of epigenetic change that occurs in *ddm1* mutant backgrounds, despite the global effect of reduced cytosine methylation, is new methylation and silencing of genes that are normally unmethylated and expressed. A well-characterized example is the *SUPERMAN* transcription factor gene, which controls flower development. In *ddm1* mutant plants, *SUPERMAN* frequently acquires methylation *de novo*, resulting in reduced expression and abnormal flower development. The underlying cause of this *de novo* methylation is not known, but an attractive possibility is that a *ddm1*-induced epigenetic change at some other region of the genome creates a signal that targets *SUPERMAN* for silencing. As in other cases of paramutation, this signal could either be transmitted by direct chromatin contacts, or it could be transmitted by a mobile intermediate such as an RNA molecule. The *de novo* methylation of *SUPERMAN* also shows that specific methylation imprints can still be established and maintained in the absence of DDM1 function.

## Epigenetics and plant genetic engineering

In the past decade there has been an increase in the modification of agricultural plants with transgene insertions that express desirable traits. However, a frequent stumbling block is the unwanted silencing of the transgene. Studies of silenced transgenes and of silenced endogenous sequences have shown that repeated sequence arrays, particularly inverted repeats, are most prone to silencing. Thus, selection of transgenic plants with single-copy transgene insertions is the first line of defense against silencing. Even with this safeguard, the transgene might still be silenced if it inserts near heterochromatin or if it expresses an RNA trigger for silencing. But with an increased understanding of epigenetic patterning across the plant genome, and the identification of novel gene products that control gene silencing, the tools are now in place to manipulate more effectively plant transgene expression with the goal of improving agriculture.

## Additional Reading

Richards, E.J. and Elgin, S.C.R. (2002). Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. Cell *108*, 489–500.

Dorweiler, J.E., Carey, C.C., Kubo, K.M., Hollick, J.B., Kermicle, J.L. and Chandler, V.L. (2000). *mediator of paramutation1* is required for establishment and maintenance of paramutation at multiple maize loci. Plant Cell *12*, 2101–2118.

Mette, M.F., Aufsatz, W., van der Winden, J., Matzke, M.A. and Matzke, A.J.M. (2000). Transcriptional silencing and promoter methylation triggered by double-stranded RNA. EMBO J. *19*, 5194–5201.

Bartee, L., Malagnac, F. and Bender, J. (2001). *Arabidopsis cmt3* chromomethylase mutations block non-CG methylation and silencing of an endogenous gene. Genes Dev. *15*, 1753–1758.

Richards, E.J. (1997). DNA methylation and plant development. Trends Genet. *8*, 319–323.

Jacobsen, S.E., Sakai, H., Finnegan, E.J., Cao, X. and Meyerowitz, E.M. (2000). Ectopic hypermethylation of flower-specific genes in *Arabidopsis.* Curr. Biol. *10*, 179–186.

Department of Biochemistry and Molecular Biology, Johns Hopkins University Bloomberg School of Public Health, 615 N. Wolfe St., Baltimore, Maryland 21205, USA.
E-mail: jbender@mail.jhmi.edu